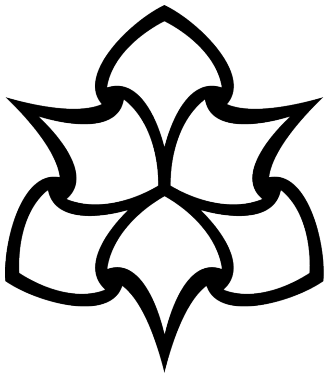

Objective Analysis of Neck Muscle Boundaries for Cervical Dystonia Using Ultrasound Imaging and Deep Learning

Ian Loram, *Member, IEEE*, Abdul Siddique, María B. Sánchez, Pete Harding, Monty Silverdale, Christopher Kobylecki, and Ryan Cunningham, *Member, IEEE*

Abstract*—*Objective: **To provide objective visualization and pattern analysis of neck muscle boundaries to inform and monitor treatment of cervical dystonia. *Methods:* We recorded transverse cervical ultrasound (US) images and whole-body motion analysis of sixty-one standing participants (35 cervical dystonia, 26 age matched controls). We manually annotated 3,272 US images sampling posture and the functional range of pitch, yaw, and roll head movements. Using previously validated methods, we used 60-fold cross validation to train, validate and test a deep neural network (U-net) to classify pixels to 13 categories (five paired neck muscles, skin, ligamentum nuchae, vertebra). For all participants for their normal standing posture, we segmented US images and classified condition (Dystonia/Control), sex and age (higher/lower) from segment boundaries. We performed an explanatory, visualization analysis of dystonia muscle-boundaries. *Results:* For all segments, agreement with manual labels was Dice Coefficient (64 $\pm$ 21%) and Hausdorff Distance (5.7 $\pm$ 4 mm). For deep muscle layers, boundaries predicted central injection sites with average precision 94 $\pm$ 3%. Using leave-one-out cross-validation, a support-vector-machine classified condition, sex, and age from predicted muscle boundaries at accuracy 70.5%, 67.2%, 52.4% respectively, exceeding classification by manual labels. From muscle boundaries, Dystonia clustered optimally into three sub-groups. These sub-groups are visualized and explained by three eigen-patterns which correlate significantly with truncal and head posture. *Conclusion:* Using US, neck muscle shape alone discriminates dystonia from healthy controls. *Significance:* Using deep learning, US imaging allows online, automated visualization, and diagnostic analysis of cervical dystonia and segmentation of individual muscles for targeted injection.**

***Index Terms*—Deep learning, ultrasound imaging, cervical dystonia, segmentation, muscle boundaries, diagnosis.**

Manuscript received April 30, 2019; revised October 31, 2019 and December 23, 2019; accepted December 30, 2019. Date of publication January 9, 2020; date of current version April 6, 2020. This work was funded by The Dystonia Society Grant A clinical tool for real-time analysis and visualization of cervical muscles for cervical dystonia, by Department of Life Sciences, Manchester Metropolitan University (MMU) for R Cunningham and by Salford Royal Foundation Trust (SRFT) Clinical Research Fellowship for A Siddique. *(Corresponding authors: Ian Loram; Ryan Cunningham.)*

I. Loram is with Cognitive Motor Function Research Group, Research Centre for Musculoskeletal Science & Sports Medicine, MMU, M1 5GD Manchester, U.K. (e-mail: i.loram@mmu.ac.uk).

M. B. Sánchez is with the Department of Health Professions, MMU (e-mail: m.sanchez.puccini@mmu.ac.uk).

R. Cunningham is with the Centre for Advanced Computational Science, Department of Health Professions, MMU (e-mail: ryan.cunningham@mmu.ac.uk).

A. Siddique is with the Department of Neurology, SRFT, M6 8HD Salford, U.K. (e-mail: abdul.siddique@srft.nhs.uk).

M. Silverdale and C. Kobylecki are with the Department of Neurology, Salford Royal NHS Foundation Trust, Manchester Academic Health Sciences Centre, The University of Manchester, Oxford Road Manchester, M13 9PT Manchester, U.K. (e-mail: monty.silverdale@manchester.ac.uk; christopher.kobylecki@manchester.ac.uk).

P. Harding is with the Elements Technology Platforms Ltd, S1 2NS Sheffield, U.K. (e-mail: pete@elementstechnology.co.uk).

Digital Object Identifier 10.1109/JBHI.2020.2964098

I. Introduction

CERVICAL Dystonia (CD), also called *spasmodic torticollis*, is a painful condition in which the neck muscles contract involuntarily, causing the head to twist, turn, and pull into an abnormal posture. This neurological movement disorder affects an estimated 18,000 adults in the UK [1]. The reported mean duration from symptom onset to diagnosis is 44 months, with consultations sought from a mean of 3.5 different healthcare providers before reaching a diagnosis and receiving effective therapy [2]. For CD, the diagnosis is based on expert clinical assessment since laboratory testing and imaging of the brain or spine is typically unrevealing [2].

Treatment of CD is symptomatic and the established protocol is injecting the neck muscles with botulinum neurotoxin (BoNT) [3]. Clinical experience shows the main causes for treatment failure are suboptimal neck muscle selection or BoNT dosing, indicating the importance of appropriate targeting of overactive muscles [4], [5]. Furthermore, monitoring the effectiveness of treatment is confounded by use of differing rating scales and assessment methods [6]. There is a clinical need to diagnose CD more promptly, to improve analysis and identification of dystonic muscles, to improve delivery of injection and dose to specific muscles, to provide objective recording of injection sites for retention within medical records and to track longitudinally the effect of injections on individual muscles [7].

A. Current Clinical Methods

The most common method of identifying and injecting the muscles involved in CD is clinical examination and manual

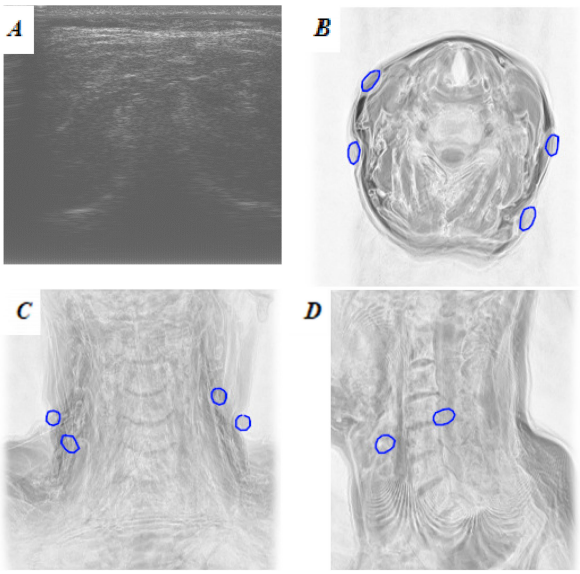

Fig. 1. **Posture (linked US-MRI) datase. (A)** Representative axial ultrasound image targeted at vertebral level C4. The probe plane was marked with four cod liver oil capsules. (B), (C), (D): Axial, coronal and sagittal MR images of the same participant showing the ultrasound image plane marked by four cod liver oil capsules (blue circles). Here, the ultrasound image plane lies between cervical vertebrae C3 and C4. Note the challenge of extracting muscle boundaries from the ultrasound image. For the whole dataset, images were acquired at level 3.8 ± 0.6 (mean ± SD).

needle placement, based on the clinician's knowledge of functional anatomy of the neck muscles, directed by head position, shoulder elevation and assessment of muscle tone and hypertrophy on palpation [8]. Within the clinic, this method may be efficacious for superficial neck muscles but is compromised for deeper muscles - typically the deeper neck muscles (e.g., spinalis cervicis, multifidus) are difficult to assess clinically and inject. There is increasing awareness of the clinical relevance of deep cervical muscles in the pathogenesis and potential therapy of CD [9], [10], but the tools to assess and treat these muscles are currently not fully developed. Compared with intramuscular electromyographic (iEMG) mapping of cervical muscle activity, the sensitivity of clinical examination has been reported as 59% and the specificity 75% [8]. The positive predictive value of shoulder elevation and muscle hypertrophy is reportedly only 70% and head position does not provide added value, because individuals with solitary dystonic head postures do not have muscle dystonia following simple patterns [8]. Without iEMG mapping, 41% of dystonic muscles would not be recognized and 25% of inactive muscles would be judged dystonic [8]. However, iEMG is time consuming, requires substantial expertise, is invasive and cannot be performed in individuals on anticoagulants. Other methods including measurement of electrical impedance have been proposed to be sensitive to muscle changes in CD, but are not as yet validated [11].

Ultrasound (US) offers non-invasive visualisation of muscle structures with easy contralateral comparison, is readily available, and improves the precision of injections [12], [13]. However, use of US requires training, is dependent upon operator expertise, and remains subjective [12].

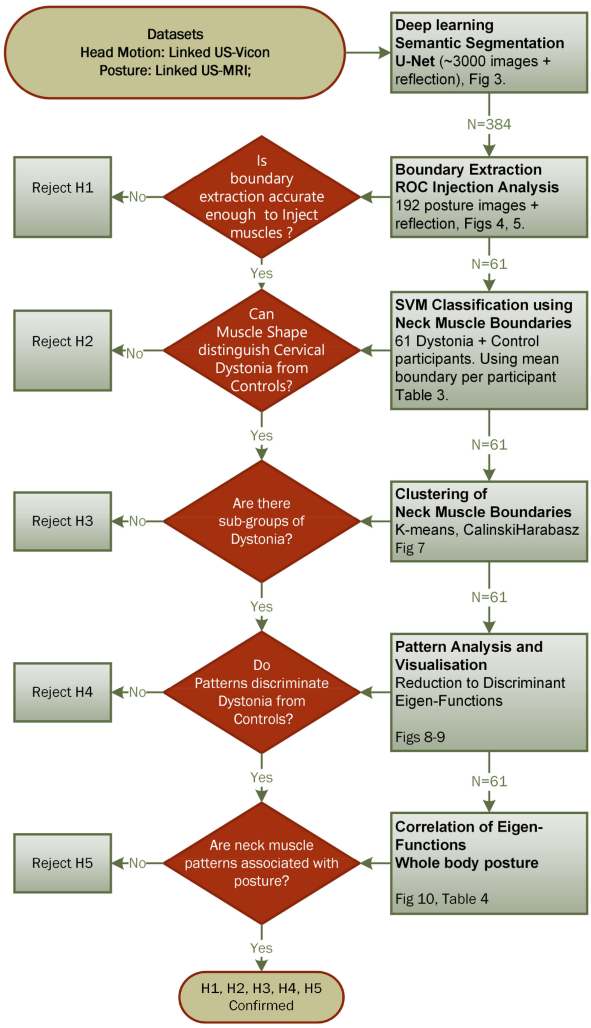

Fig. 2. **Hypothesis pipeline:** Using Deep Learning Semantic Segmentation, Boundary Extraction, support vector machine (SVM) Classification, Clustering, Pattern Analysis, Visualisation and whole body motion analysis, we sequentially test five hypotheses concerning transverse ultrasound images of the human neck. **H1**, Segmentation is accurate enough to guide injection of deep neck muscles (Figs. 4, 5). **H2**, Cervical Dystonia can be classified from age matched controls using muscle shape alone (Table III). **H3**, Muscle shape clusters into subtypes of Cervical Dystonia (Fig. 5). **H4**, Cervical Dystonia can be reduced to significant eigen-patterns of muscle shape (Figs. 6–8). **H5,** Eigen-patterns of neck muscle shape are associated with features of whole body posture (Fig. 8).

The objective of this study is to provide an automated, objective visualization of neck muscle boundaries and to analyze whether these boundaries have diagnostic value discriminating patterns of cervical dystonia from healthy controls. If successful, these methods demonstrate proof of concept for a clinical tool for objective online diagnosis, injection guidance and monitoring, with minimal requirement for operator expertise and minimal burden on clinical time.

B. Contribution of This Study

The use of deep learning to extract information from limited quality images (c.f. Fig. 1A) is progressing rapidly. Application

to US is under-developed and application to skeletal muscle is rare. This study builds upon previous work by our group realizing the scientific and clinical value of *in-vivo* skeletal muscle analysis [14]–[19], applying deep learning to skeletal muscle US [20]–[22] and specifically developing methods for analysis of the neck muscles [23], [24]. Recently we contributed a dataset, a methodology for labelling training images suitable for participants with involuntary head movement, and a benchmark deep learning method for segmenting the neck muscles [24], [25]. Here we apply that methodology to cervical dystonia.

Our primary interest is to establish whether from a single axial image, neck muscle shape allows differentiation of cervical dystonia from healthy controls. Our secondary interest is whether or not segmentation is accurate enough to guide injections to deep neck muscles.

Cervical dystonia is characterized by the sustained contraction of specific muscles. Our general hypothesis is that those contracting muscles cause an identifiable pattern of neck muscle shape and an associated pattern of whole body posture. Dystonic muscles are specific to an individual, but common combinations are observed. Each dystonic combination should produce a pattern of neck muscle shape or posture away from the normal distribution, along a dimension which is distinct from other combinations.

We test in sequence five specific hypotheses (Fig. 2): (i) Segmentation will identify injection points within deep neck muscles accurately. (ii) Cervical dystonia can be classified from age matched controls using neck muscle shape alone. (iii) Dystonia clusters into natural sub-groups using neck muscle shape (iv) Dystonic muscle shapes can be expressed as significant eigen-patterns, (v) Dystonic muscle eigen-patterns are associated with patterns of whole body posture.

II. Methods

A. Data Collection

Using a probe (7.5 MHz, SonixTouch, Ultrasonix, USA) held transversely to the posterior neck, B mode US images (depth 5cm), were recorded from 61 adults: 35 cervical dystonia (mean age 61 $\pm$ 10, 15 male) and 26 age matched controls, (mean age 59 $\pm$ 14 years, 18 male) while standing and while performing head rotation tasks defining their range of pitch, yaw and head rotation. Power and contrast were adjusted per participant using visual feedback. We disabled image enhancement processes to reduce internal frame averaging. Images were interpolated (bi-linear) to a size, and resolution common to our previous datasets (491 $\times$ 525 pixel, 10 pixels per mm) [24]. These experiments, performed in the Faculty of Science and Engineering, Manchester Metropolitan University (MMU), received ethical approval from the NHS Health Research Authority (REC: 15/NW/0016, IRAS:169803) and from MMU Science and Engineering Faculty Ethics Committee. The study was conducted in accordance with the Declaration of Helsinki guidelines. All values are reported as mean $\pm$ SD unless stated otherwise.

***Posture** (Linked US-MRI) Dataset:* These posture images are the subject of this paper. Participants stood upright, observing a monitor at 1m distance, just below eye level. Three or more

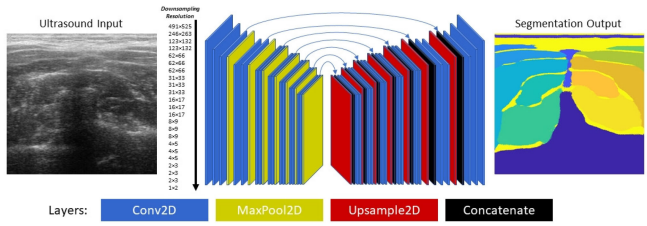

Fig. 3. **U-Net Model Architecture.** This figure details the best performing model, according to Table I Supplementary Material. The model consists of 2D convolutional and pooling layers in the encoder part of the network (blocks to the left), and 2D up-sampling, concatenation and convolutional layers in the decoder part of the network (blocks to the right), where concatenation layers concatenate up-sampled layers along the feature channels with compatible layers in the encoder network enabling flow of information and gradients in forward and backward passes, respectively. This neural network has over 51,000,000 trainable parameters, and over 21,000,000 functional outputs, and operates in real-time (*approx.* 10 frames per second) on a modest PC or laptop. Hyperparameters and augmentation were fixed, chosen based on the experience and well-established literature. **Hyperparameters**: *Dropout* = 0.25,L^2 = 0.0005, *Adam*(α = 0.00005, β_1 = 0.9, β_2 = 0.999), *BatchSize* = 1,*Epochs* = 40. **Data Augmentation**: *Local Contrast Normalisation* = 31 $\times$ 31, *Rotation* = $\pm$ 8°, $Trans_x = \pm$ 128, $Trans_y = \pm$ 64

axial US images of the posterior neck targeted at level C4 were recorded, each with renewed probe placement (Fig. 1A). Four cod liver oil capsules were taped (using Transpore medical tape) to the neck, two either side of the neck approximately in the image plane of the probe. The probe was removed, leaving the capsules in place, and an MRI scan (0.3T open MRI scanner, G-Scan, Esaote, Italy) was obtained with participants lying supine on the scanning bed and their neck positioned central within a cervical imaging coil. Axial scans (Spin T1-weighted HF, matrix 512 $\times$ 512) were performed in a range from the upper jaw line to the clavicle, orthogonal to the spine, in 19 equidistant sections (Fig. 1B-D). Manual labelling of muscle boundaries in US images is challenging. The purpose of collecting linked MRI-US data is to train experts in labelling US images using methods reported previously [24].

***Head Motion** (Linked US-Vicon) Dataset:* In a separate session, forty seven retroreflective markers were attached to the body to allow motion analysis of eighteen body segments (head, neck, thorax, pelvis, thighs, shanks, feet, clavicles, upper arms, forearms, hands).

Participants stood in the middle of the calibrated volume and were instructed to perform pitch (flexion/extension), yaw (right/left) and roll (right side/left side) head rotations, turning their heads as far as possible in both directions. Each trial was repeated starting in the opposite direction. Body motion was recorded by a 9 camera Vicon MX motion capture system. For each trial, the US probe was held to the posterior neck targeted at level C4, to allow free movement of the head and image of 5 bilateral layers of muscles. Images were saved digitally at 10 Hz with start time synchronized to the Vicon recording.

The purpose of collecting linked Vicon-US data is to investigate the relationship between neck muscle boundaries and posture/movement. Whole body kinematic data provides an additional modality of explanation and validation of the information content of neck muscle boundaries in US images

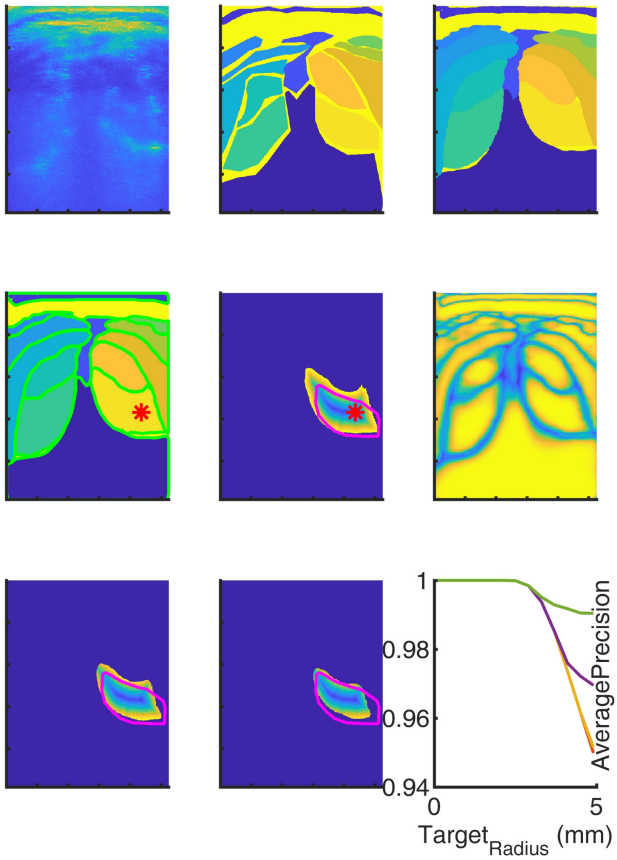

Fig. 4. **Boundary Extraction and Injection Point Analysis.** This illustrates the output and visualisation provided by the neural network approach. This participant had cervical dystonia (note the asymmetry). **Top row**: **Left**: ultrasound image. **Middle**: Manually defined labels. **Right**: neural network predicted classification of pixels. **2nd row**: **Left:** boundaries (green) estimated around each segment. **Middle**: Injection point (red star) defined as pixel most distant from predicted segment boundary. Colour spectrum blue to yellow shows decreasing distance from predicted boundary and hence increasing target margin around injection point. Boundary of right multifidus label (green). **3rd row**: **Left and middle**: as 2nd row middle, but predicted pixels shown restricted to confidence greater than 60 and greater than 80% respectively. **Right**: confidence of the neural network (yellow = high confidence) in classifying pixels. **Right**: Average precision of pixels within target region for right multifidus. Lines red, yellow, magenta, green show pixels restricted to confidence greater than 20, 40, 60, 80% respectively. The clear visualisation of predicted segment (top right) will be appreciated by clinicians. The visualisation of confidence of classifying pixels (2nd row right) gives the user feedback regarding optimal probe location and orientation.

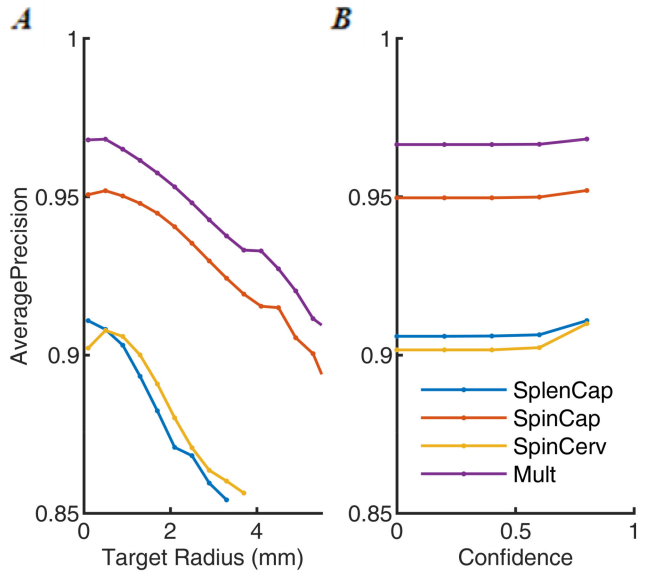

Fig. 5. **Precision of Injection points.** Testing Mode results: (A) Average Precision for all pixels of confidence more than 80% in the target region for a variety of sizes of target region. (B) Average Precision for varying levels of confidence. Precision is number accurate as a percentage of pixels predicted to be within muscles. Average precision is precision averaged with respect to recall, sorted in descending confidence.

B. Image Labelling

Using published methods two annotators were trained to a common standard using MRI images linked to US images [24]; their agreement is shown Supplementary Material (SM) (SM-Table IV). US images (192 total, ∼3 per participant) from the Posture (linked US-MRI) dataset were labelled manually by annotating the boundaries around ten muscles, vertebra, ligamentum nuchae and skin. As described previously [23], [24], MRI images showing the same cod liver oil capsule marked plane were annotated and registered to the US images to guide annotation of the US images.

Two thousand US images (∼30 per participant) from the Head Motion (linked US-Vicon) dataset, sampling uniformly the range of pitch, yaw and roll head rotations [24] were labelled manually for the same 13 segments.

Image-labels from this Cervical Dystonia Project (CDP) were supplemented by our previous posture dataset (25 linked US-MRI neck image-labels) [23] and our previous Head Motion dataset (1100 linked US-Vicon neck image-labels) [24]. These supplementary image labels were acquired using a different US probe and machine (7.5 MHz T shaped probe, taped to the neck, Aloka) at the same neck location and for the same posture and head motion tasks.

C. Machine Learning

We divided data into independent training, validation and test datasets. We report a 'Testing Mode', and an 'Analysis Mode' of division. 'Testing Mode' is used for testing Segmentation, Boundary extraction and SVM classification of dystonia, sex and age (Fig. 2).

Testing Mode: 60-fold leave-one-out (LOO) cross validation. The 61 CDP participants (2192 image-labels) were assigned into 60 folds (one per fold, except one fold contained two participants). These 60 folds provided 30 groups each containing a 'test' (one participant), 'validation' (one participant) and 'train' (58 participants) dataset. Each 'train' dataset was supplemented by the 1100 Head Motion image-labels [24].

Analysis Mode: All 3100 (2000 + 1100) Head Motion image-labels were assigned to the 'train' dataset. All 217 (192 + 25) Posture image-labels were assigned to the

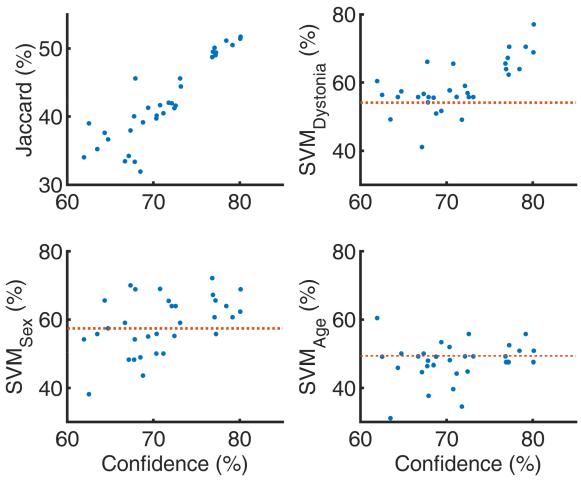

Fig. 6. **Classifying Dystonia, sex and age from segment boundaries.** For 36 deep neural networks of varying architecture and hyper parameters dating from the very start of this investigation tested on the Posture dataset (61 participants), panels show Jaccard Index, and SVM classification accuracy for Dystonia, sex and age v mean predicted confidence. Points of confidence <75%, are early 10 fold cross validation networks using only the Posture dataset (∼200 image-labels). Points of confidence >75% include addition of Head Motion dataset (∼3000 image-labels) to neural network trained in Analysis Mode, except one point which is the Testing Mode U-net. Red horizontal lines show SVM classification using the manually annotated boundaries used to train the neural networks. **Key Points**: Neural networks become better at extracting information than the human annotated ground truth used to train them. Ultrasound neck images contain the information required to classify Dystonia. Sex information is contained less well and age is not revealed within these US images. Jaccard Index increases with confidence. Classification of condition, sex or age requires US information content in addition to accurate segmentation.

'validation' and 'test' datasets and participants were assigned alternately to the 'validation' and 'test' datasets.

The 'Testing Mode' maintains strict independence between training, validation and test dataset since participants do not overlap folds. To maximize the training set in each fold, we performed LOO cross-validation. Since, some participants produce poorer quality images (e.g., deep fat layer, indistinct muscles), validation (selection of the training iteration to use for testing) will be sub-optimal. Typically, validation scores over fit prematurely before the neural network fully encodes the generalizable content of the data. 'Testing Mode' represents the harshest possible testing regime.

'Analysis Mode' maximizes the training set (all Head Motion data), and also maximizes the number of participants (all Posture data) in each validation and test set. We propose motion of the head, (and to a lesser extent repositioning the probe to a new location, pressure and orientation), produces more independence between images than changing participant. Moving the head changes the depth, muscle shape, scale, texture and dropout of each image. We propose, 'Analysis Mode' maintains independence between training (Head Motion) and validation/test (Posture) sets, and allows learning to extract more fully the content of the data.

Manual annotators were blind to the condition and sex of the image and these labels played no part in training, validation or testing. To select the mode with best descriptive power for Pattern Analysis and Visualisation (Fig. 2), we selected the mode ('Testing' v 'Analysis') with the highest SVM classification of dystonia and sex.

Augmentation: Each US image and corresponding label was flipped about the vertical line of symmetry, to double each training, validation and test set and to remove asymmetry bias from each process of training, validation and testing.

Implementation: Following previous work [24], and using software written within this group, we conducted extensive evaluation of 99 trained neural networks (c.f. SM for detail). The best encode-decoder neural-network (U-Net) was trained (Fig. 3). Training error between labels and network prediction was computed using a class-weighted cross-entropy cost function $-\sum_{i=0}^{m}\beta^{(k)}y_{(k)}^{(i)}\log\hat{y}^{(i)}$ where m is the number of pixels in a single image, i is the index of a pixel in a single image, k is the class associated with the pixel i, $y_{(k)}^{(i)}$ is the label category (0 or 1), $\hat{y}^{(i)}$ is the SoftMax response, and $\beta^{(k)}$, which up-weights $(\beta^{(k)} \geq 1)$ a given class using, $\beta^{(k)}=\frac{\max(C)}{C^k}$, where C^k is the total count of pixels of class k, and $\max(C)$ is the total count of pixels of the class with the maximum total pixel count.

Network training consisted of online learning, interrupted every quarter pass (550 learning iterations) through the training set, to record cross entropy test results from the validation and test (test) sets. If the cross-entropy loss for either test set was lower than any previous recorded loss for that test set, the network was saved to long term storage. Each selected network was tested by the other set, and *vice versa* for both networks. This process yielded held-out test results for all images in both test sets. Training terminated after 35 epochs.

The Posture dataset and predicted output was used for post neural network analysis. Please refer to Fig. 2 which defines the flow of hypothesis, methods and results.

D. Boundary Extraction

The Posture dataset was used for boundary extraction and injection point analysis (Figs. 4, 5). To the classified pixels (Fig. 4) we applied an 8×8 pixel median filter, filled holes, smoothed the boundaries and extracted boundaries using MATLAB functions (*medfilt2, imfill, imclose* and *bwtraceboundary* respectively. All boundaries were extracted clockwise, starting from a key point, defined as the most medial pixel for muscles, and interpolated to 100 evenly spaced points (Fig. 3). For one image, the pattern of 13 segments is described by a row vector or 2,600 numbers (100 horizontal, then 100 vertical coordinates for each segment). Accuracy of extracted boundaries was assessed using Jaccard Index (JI), Dice Coefficient (DC), Hausdorff Distance (HD) and modified Hausdorff Distance (MHD) (Table I) [25].

For each segment, the central predicted injection point was defined as the pixel of maximum distance ($d_{\max}$) from any boundary point (Fig. 3, middle row). We iteratively increased the margin around this injection point by distance $t = 0, 1, 2 \ldots d_{\max}$ mm. The pixels enclosed by this boundary at distance $d_{\max} - t$ from the predicted segment boundary provided a series of target injection regions (Fig. 4). By comparison with

TABLE I

BOUNDARY ACCURACY BETWEEN BOUNDARIES EXTRACTED FROM NEURAL NETWORK OUTPUT AND MANUAL ANNOTATION. JACCARD INDEX (JI) AND DICE COEFFICIENT (DC) SHOW PERCENTAGE INTERSECTION OVER UNION. HAUSDORFF DISTANCE (HD, IN mm) SHOWS THE GREATEST DISTANCE, AND MODIFIED HAUSDORFF DISTANCE (MHD, IN mm) SHOWS THE MEAN DIFFERENCE BETWEEN PREDICTED AND MANUALLY ANNOTATED BOUNDARIES. THIS TABLE REPORTS MEAN ± S.D. FOR ALL IMAGES FROM THE POSTURE DATASET (N = 384 = 192 + 192 REFLECTIONS) USING 'TESTING MODE'. THE PENULTIMATE COLUMN ALONE REPORTS ALL SEGMENTS USING 'ANALYSIS MODE'. THE FINAL COLUMN REPORTS AGREEMENT BETWEEN THE TWO ANNOTATORS WHO CONTRIBUTED THE LABELS FOR THIS DATASET

	TESTING MODE									ANALYSIS MODE	Inter Expert
Metric	Vertebra	Multifidus	Spinalis. Cervicis	Spinalis Capitis	Splenius Capitis	Trapezius	Lig. Nuchae	Skin	All Segments	All Segments	All Segments
JI (%)	78±12	54±12	51±18	54±14	50±16	37±17	35±15	47±17	50±21	51±21	46±20
DC (%)	87±8	68±12	65±19	68±13	64±17	51±19	50±18	62±16	64±21	64±21	59±23
HD (mm)	5.8±2.5	6.6±2.4	5.1±2.0	6.0±2.4	6.3±3.5	5.4±2.8	7.2±4.7	2.3±1.5	5.7±3.7	5.6±3.6	6.0±4.3
MHD (mm)	1.4±0.8	2.5±1.0	2.0±1.0	1.6±0.7	1.7±1.0	1.6±1.0	2.5±3.2	0.8±0.3	1.8±1.5	1.8±1.4	2.0±2.3

corresponding pixels in the manually labeled image, we compute average precision, for varying target region. This analysis was iterated using pixels only of predicted confidence (SoftMax scores) greater than 0, 0.2, 0.4, 0.6 and 0.8 respectively (Figs. 4, 5).

E. SVM Classification Using Boundaries

Boundaries from reflected images were discarded. For all 61 participants we computed the mean segment boundaries. This generated a matrix of 61 rows by 2600 columns. With Matlab functions *fitcsvm*, *crossval* and *kfoldLoss*, we used a support vector machine, with 61-fold LOO cross validation to test ability of the boundaries to classify clinical condition (Dystonia v Control), Sex (male v female), and age (higher, lower) where age as divided into two groups around the median value (Table III).

F. Clustering of Boundaries

Using k-means, we tested the extent to which the 61 × 2600 matrix of segment boundaries clustered into groups. We used Matlab function *kmeans* with correlation as the distance metric. We used the CalinskiHarabasz value to evaluate separation into 2 to 10 groups. (Fig. 7).

G. Pattern Analysis and Visualisation

In a series of steps, we reduced the boundaries to the statistically significant eigen-patterns which discriminate dystonia sub-groups and healthy controls (Figs. 8, 9).

Using all images from the Posture dataset, we reduced 2600 columns to 100 principal components. Each component represents a pattern of variation from the mean shape. Using unreflected cases only, we computed the mean principal component scores for the dystonia and control participants (n = 61). We selected the principal components which reconstruct the group membership (Dystonia 1-3, Control). To select, we computed a univariate ANOVA for each principal component. Then, using MATLAB functions *sequentialfs* (10-fold cross validation, 50 monte-Carlo repetitions, forward entry starting with significant univariate components), and *classify* ('diaglinear', naive Bayes), we selected the combination of principal components which predicts group membership.

To reduce the model to statistically significant discriminant eigen-functions, maximizing separation of the groups, we performed one-way Multivariate Analysis of Variance (n = 61), using MATLAB function *manova1* (Figs. 8, 9);

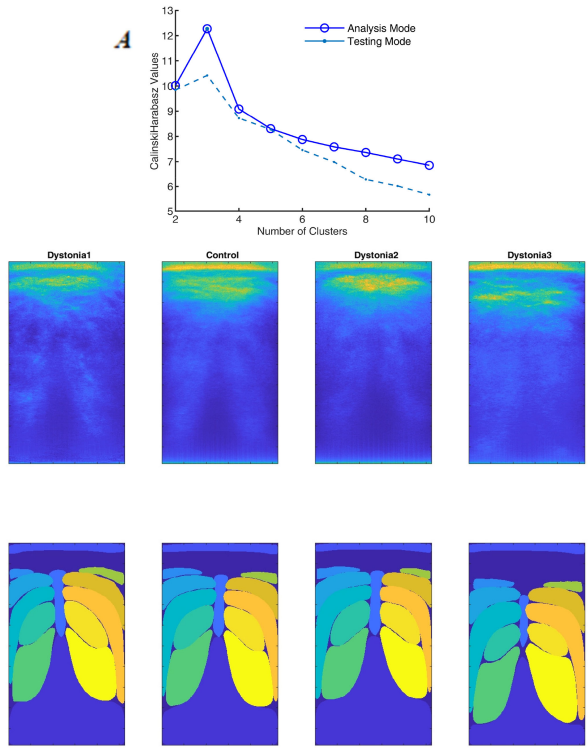

Fig. 7. **Clustering Dystonia into sub-groups** Dystonia participants were clusterred into sub-groups using segment boundaries and k-means algorithm. (A) Clustering success metric v numer of clusters. (B) Division of Dystonia into optimal number of sub-groups (Dystonia 1-3, n = 9, 17, 9 respectively) using Analysis Mode. For each group we show: **Top Row**. Group averaged ultrasound images, **Bottom Row**. Mean segment bounadaries. Left side of all images represents the left anatomical side of the participants. Dystonia groups 1-3 appear left side compressed, right side compressed and both sides compressed respectively.

TABLE II

MUSCLE INJECTION ACCURACY. SHOWS CLASSIFICATION ACCURACY FOR PIXELS OF CONFIDENCE GREATER THAN 80% WITHIN TARGET REGION. VALUES ARE MEAN FOR ALL PARTICIPANTS IN TEST SET (N = 61). DEEP MUSCLES SHOWS MEAN ± S.D. FOR MULTIFIDUS TO SPLENIUS. INJECTION POINT IS THE PIXEL FURTHEST FROM ALL BOUNDARY POINTS. TARGET RADIUS: DECREASE IN DISTANCE FROM EDGE TO TARGET REGION INCREASE TARGET AREA. DISTANCE FROM EDGE: DISTANCE OF TARGET BOUNDARY FROM SEGMENT BOUNDARY. N POS, N NEG: NUMBER OF PIXELS IN IMAGE. ACC: PERCENTAGE OF PIXELS CLASSIFIED CORRECTLY. TP RATE, FP RATE, FN RATE, TN RATE: TRUE POSITIVE, FALSE POSITIVE, FALSE NEGATIVE, AND TRUE NEGATIVE RATE RESPECTIVELY. AP: AVERAGE PRECISION (PRECISION AVERAGED WITH RESPECT TO RECALL SORTED IN DECREASING CONFIDENCE)

Segment	Target Radius (mm)	Distance From Edge (mm)	N Pos ($x10^6$)	N Neg ($x10^6$)	Acc (%)	TP Rate (%)	FN Rate (%)	FP Rate (%)	TN Rate (%)	AP (%)
Multifidus	0.3	5.9	4.7	61.5	93.1	1.5	98.5	0.0	100.0	97.2
Spinalis Cervicis	0.3	3.7	2.6	63.5	96.1	3.2	96.8	0.0	100.0	94.8
Spinalis Capitis	0.7	2.8	3.4	62.8	95.4	12.5	87.5	0.2	99.8	89.9
Splenius Capitis	0.3	2.4	2.2	63.9	96.7	3.2	96.8	0.0	100.0	92.2
Trapezius	1.1	0.1	0.7	65.5	99.2	55.4	44.6	0.3	99.7	82.8
Deep Muscles	**0.4 ±0.2**	**3.7 ±1.6**	**3.2 ±1**	**62.9 ±1**	**95.3 ±1.6**	**5.1 ±5**	**94.9 ±5**	**0.01 ±0.01**	**99.9 ±0.01**	**93.5 ±3**

H. Correlation of Eigen-Functions With Whole Body Posture

For each participant, we computed their median multi-segment posture (51 angular components from 17 joints) from all their trials in the Head Motion (*linked US-Vicon*) dataset. To identify joint angles associated with neck boundary eigen-functions, we calculated the structure matrix showing correlation (n = 61) of all joint angles with each eigen function and limited the lists to those significant at $p < 0.05$ (Table IV).

III. RESULTS

Supplementary Material, presents comparative analysis of 96 trained neural networks from ten different models extending our previous work [24] and justifying the best model selected for this paper. The video in SM demonstrates live neural network output of the selected model.

Figure 2 defines the flow of hypotheses and results presented below. We report five main findings: (i) accuracy of extracted boundaries and of injection points within neck muscles, (ii) classification of condition, sex and age from muscle boundaries, (iii) the optimal clustering of dystonia into sub-groups, (iv) reduction to eigen-patterns of muscle shape associated with cervical dystonia and (v) the association of neck muscle eigen-patterns with whole body posture.

A. Accuracy of Extracted Boundaries and Injection Points

For the Posture dataset, accuracy of all segment boundaries using metrics JI, DC, HD and MHD was equal using '*Testing Mode*', or using '*Analysis Mode*' and both marginally higher than inter-annotator agreement (Table I). These values were typical for muscles deep to the surface (Multifidus, Spinalis Cervicis, Spinalis Capitis, Splenius Capitis), (Table I).

A meaningful assessment of accuracy is provided by the question "would an injection into the predicted segment target the proposed muscle accurately?" Predicted classification of pixels is more confident towards the center of the muscles rather than at the boundary (Fig. 4). By using SoftMax confidence at 80% or more to select injection points (Fig. 4), accuracy indicated by average precision is improved (Fig. 5). Setting minimum prediction confidence to 80%, accuracy of injection points for the deep muscles (multifidus, spinalis cervicis, spinalis capitis, splenius capitis) is indicated by average precision 93.5 ± 3% (Table II). These results support our hypothesis (i) that segmentation of deep muscles will identify injection points within the designated muscle.

TABLE III

CLASSIFICATION OF DYSTONIA, SEX AND AGE FROM MUSCLE BOUNDARIES. A SUPPORT VECTOR MACHINE, WITH LOO 60 FOLD CROSS VALIDATION PREDICTED CONDITION (DYSTONIA V CONTROL), SEX (MALE VS. FEMALE) AND AGE (ABOVE MEDIAN VS. BELOW MEDIAN) FROM POSTURE DATASET OF 61 PARTICIPANTS

Source	Condition	Sex	Age
Testing Mode Predicted boundaries	70.4%	67.2%	52.4%
Analysis Mode Predicted boundaries	77.0%	68.9%	50.8%
Manually annotated boundaries	54.0%	57.%	49.2%
Joint Angles	63.9%	70.4%	49.2%

B. SVM Classification Using Segment Boundaries

Using Testing Mode results, and predicted neck segment boundaries alone as input, a support vector machine with LOO cross validation, classified Dystonia from age matched controls with accuracy 70.4%, which was higher than sex (67.2%) or classification age (52.4%) (Table III). This classification was higher than classification from manual annotated boundaries at 54%, 57% and 49% for condition, sex and age respectively. Classification of Dystonia and sex from boundaries was higher using Analysis Mode (77.0%, 68.9% respectively), than using Testing Mode (Table III). Results since the start of our investigation have been consistent: 'Dystonia' manifests more clearly in neck US images than sex; age cannot be classified from these images; and neural networks out-perform the human annotators (Fig. 6).

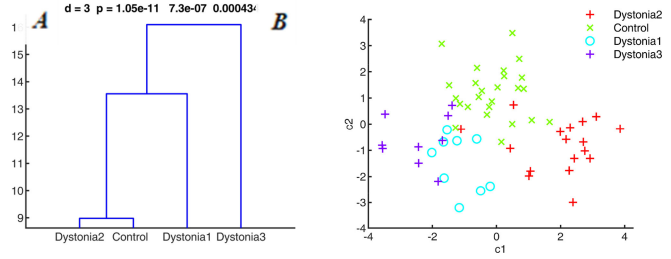

Fig. 8. **Reconstruction of Dystonia sub-groups using eigen-patterns of neck muscle shape. W**e have reduced description of dystonia sub-groups and controls to three discriminant eigen-functions (DF1, DF2, DF3, Wilk's lambda, p-values shown panel A). Each eigen-function represents a weighted combination of principal components of US muscle boundaries from all participants (35 dystonia, 26 age matched controls). N.B. "Muscle" shape refers to all 13 segments (muscle, vertebra, ligament, skin). (A) Distance (mean linkage) between group centres using Mahalanobis distance (i.e., units of within group variance, so 16 = 4 S.D. of within group variation). (B) Axes shown first two canonical discriminant function scores (c1, c2 for DF1, DF2 respectively, n = 61). Dystonia sub-groups differ from healthy controls in directions which are distinct from each other. Each eigen-function represents a pattern (Fig. 9).

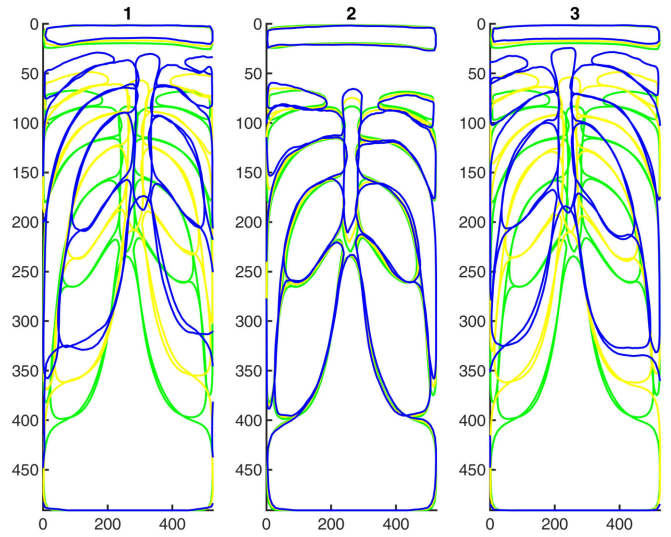

Fig. 9. **Reconstruction of Dystonia using patterns of neck muscle shape.** Sub-types of dystonia can be reconstructed from three significant eigen-patterns of neck muscle shape shown in order of significance. Green is sample mean principal component. Yellow and Blue show respectively +0.5 and +1 standard deviation of the eigen function. Image left shows anatomical left. **DF1** shows the right muscles of the neck compressed and all muscles compressed to the skin. **DF2** shows enlargement, fattening of left splenius. **DF3** shows a squashing of all muscles to the skin, more compression on left side, but more symmetrical than DF1.

C. Does Dystonia Cluster Naturally Into Sub-Groups?

Using neck segment boundaries alone, the 35 Dystonia participants clustered optimally into three groups (Fig. 7A), using both Analysis Mode and Testing Mode. Given higher SVM classification, Analysis Mode was selected for descriptive clustering and pattern analysis. The mean boundaries of these groups show 'right' asymmetry (Dystonia 2), 'left' asymmetry (Dystonia1) and 'deep' segments with a large gap between skin and muscle (Dystonia 3), (Fig. 7). A large gap could represent a thicker fat layer or could represent altered pitch of the head. The right/left asymmetry could result from a tilted head (roll), a turned head (yaw), an elevated shoulder, a laterally shifted neck or a combination.

TABLE IV

POSTURE ASSOCIATED WITH NECK EIGEN-PATTERNS (STRUCTURE MATRIX). SHOWS JOINT ANGLES ASSOCIATED WITH THREE DISCRIMINANT EIGEN-FUNCTIONS. DF1 ASSOCIATES WITH THE WHOLE BODY LEANING RIGHTWARDS, THE HEAD TILTING DOWN AND TO THE LEFT, THE LEFT SHOULDER RAISED THE RIGHT WRIST COCKED. DF2 ASSOCIATES WITH THE WHOLE BODY LEARNING RIGHTWARDS. DF3 ASSOCIATES WITH RIGHT ARM OUT, THE HEAD TILTED TO THE RIGHT, THE LEFT KNEE BUCKLED, AND THE RIGHT FOOT ROLLED OUT

	Joint angle	Function	r	p-value
DF1	RAnkleAngley	Whole body deviation rightwards	0.5	0.00004
	AOAnglex	Head looking downwards (pitch/extension)	-0.04	0.0038
	LClavicleAngley	Left shoulder elevation	0.3	0.02
	AOAnglez	Head turning leftwards (yaw)	0.26	0.042
	RWristAnglex	Right wrist curled (extension)	-0.25	0.048
DF2	LAnkleAngley	Whole body deviation leftwards	0.34	0.0076
DF3	RShoulderClavicleAngley	Right upper arm elevation (abduction)	0.37	0.0035
	AOAngley	Head tilting rightwards (roll)	0.36	0.0045
	LKneeAngley	Left knee inward	-0.29	0.024
	RAnkleAnglez	Right foot rolled (supination)	0.27	0.039

D. Can Dystonia be Reduced to Statistically Significant Eigen-Patterns of Neck Muscle Boundaries?

Four groups can be discriminated by a maximum of 3 eigen-functions. Having reduced the segment boundaries to 100 principal components, the feature selection procedure selected 15 components to predict group membership (Dystonia 1-3, Control) robustly and reconstructed group membership correctly at 85.2% using LOO classification.

Discriminant function analysis of the 61 case × 15 component matrix revealed three significant eigen-functions (DF1, DF2, DF2, $p = 1.1 \times 10^{-11}$, $p = 7.3 \times 10^{-7}$, $p = 0.0004$, Fig. 8). Fig. 8A shows that with a separation of 16 units of mahalanobis distance (i.e., 4 S.D of within group variance), Dystonia3 ('deep') differs substantially from Dystonia 1-2 and Control. At 13.5 units, Dystonia 1 'left' differs substantially from Dystonia2 and Control whereas Dystonia2 'right' differs from Control by only 8 units (Fig. 8A). Thus we expect 'deep' to represent the largest pattern of difference from controls.

The three discriminant functions (DF1-3) represent the significant dimensions separating the groups (Fig. 8B) and those functions also represent patterns of altered neck muscle shape distinguishing the groups (Fig. 9).

The first dimension (DF1) provides an axis separating Dystonia3-'deep' from Dystonia 2-'right', with Dystonia1 and Controls in the middle (Fig. 8B). DF1 as a patterns shows superficial movement of all structures, and rightwards displacement of the midline and a relative depth-wise compression of the right

muscles (Fig. 9B). Correlation of muscle areas (calculated as percentage area of all segments) with DF1 at ($p < 0.05$) shows reduced area of right Splenius ($r = -0.47$, $p = 0.00014$), right spinalis capitis ($r = -0.47$, 0.00015), right trapezius ($r = -0.43$, $p = 0.00046$) and increased area of left splenius ($r = 0.41$, $p = 0.001$), left spinalis capitis ($r = 0.3$, $p = 0.002$) and left spinalis cervicis ($r = 0.28$, $p = 0.027$).

The second dimension (DF2) provides an axis proceeding from all Dystonia groups (negative) to controls (positive) (Fig. 8B). As a pattern, DF2 shows asymmetric enlargement of the left muscles and compression of the right muscles, but no general superficial movement or sideways displacement of the midline (Fig. 9). DF2 is associated with reduced area of right splenius capitis ($r = -0.34$, $p = 0.008$). DF3 shows superficial shift of all structures, leftward displacement of the midline (Fig. 9) and is associated with significant reduction of left muscles and enlargement of right muscles.

E. Correspondence Between Neck Muscle Patterns and Whole-Body Posture

The Dystonia sub-groups, clustered from neck muscle boundaries alone are associated with patterns of whole body standing posture. The median standing postures of these groups (Dystonia 1-3, Control) show differences in whole body truncal alignment and head turn which are described in the legend. Univariate ANOVA of joint angles confirms the most significant difference between groups lies in whole body frontal lean to the right (Right AnkleAngle_y, $p = 0.0005$), left head tilt (AtlantoOccipitalAngle_y, $p = 0.003$) and head pitch (AtlantoOccipitalAngle_x, $p = 0.01$).

The discriminant neck muscle eigen-patterns associate with postural joint rotations. The associations with posture give validation and explanation to the neck muscle eigen-patterns. As shown in Table IV, DF1 is associated with whole body lean rightwards, head looking downwards (pitch/extension), left shoulder elevation, head turning leftwards (yaw) and right wrist curled (extension). DF2 is associated with whole body lean leftward and has little association with head-neck rotations. DF3 is associated with right upper arm elevation (abduction), head tilt rightwards (roll), left knee inward and right foot roll (supination).

This whole-body motion analysis, provides validation that dystonic patterns of neck muscles identified from US images, have functional correlates in the standing posture.

F. Comparison of Diagnosis by US With Diagnosis by Posture

Classification of clinical condition (dystonia v control) by neck muscle boundaries was superior to classification by standing posture (Table III). Using the SVM, with LOO cross validation ($n = 61$), we provide a comparative classification of condition, sex and age using whole body motion data. From the 51 components of joint rotation (17 joints $\times$ 3 degrees of rotation), the SVM classified condition with lower accuracy than the US based classification, sex with higher accuracy and could not predict age (Table III). These results confirm the US images of the neck provide a better basis for diagnosing cervical dystonia than body posture.

IV. Discussion

A. Contribution of This Study: The Main Results

This study reports the first application of deep learning to the segmentation, analysis and visualization of axial neck US images to participants with cervical dystonia. From a sample of 35 participants with cervical dystonia, and 26 age matched controls, we classified image pixels, extracted neck muscle boundaries, and tested ability to classify Dystonia, sex or age from muscle boundaries. We further clustered dystonia participants into sub-groups (Dystonia 1-3) identified the significant eigen-patterns, reconstructing dystonia and related those eigen-patterns to posture.

The most salient findings are: -

i) Cervical dystonia can be discriminated from age matched healthy controls, using an axial US image of the neck muscles. Leave-one-out classification of Dystonia v Control using SVM was correct at 70% (Table III)

ii) Cervical dystonia is associated with visible, explainable patterns of neck muscle shape (Fig. 7, Fig. 9). This sample showed optimally three dystonia sub-groups, resulting in three significant eigen-patterns of neck muscle shape (Figs. 7, 9).

iii) Each pattern showed characteristic changes in muscle depth, midline asymmetry-curvature and left-right muscle imbalance (Fig. 7, 9). The first (DF1) is associated with a postural pattern of head pitch, head turn, shoulder elevation and truncal tilt (Fig. 10, Table IV). The second (DF2), associated most strongly with truncal tilt. The third (DF3) associated most strongly with head tilt.

iv) Segmentation is accurate enough to guide injections to specific muscles (Table II).

v) Supervised deep learning of US muscle images, can encode information with a veracity exceeding the manual annotation of its human supervisors (Table III).

B. Rationale for Methods of Analysis

This focus of this paper is primarily scientific. In other words, do transverse ultrasound images of the neck, obtained from an ordinary ultrasound machine, contain the information necessary to inform understanding and diagnosis of dystonia and to aid delivery and monitoring of treatment by botulinum toxin injections? The development of methods for annotating images, training neural networks and evaluating deep learning architectures to segment muscle boundaries in US images of the posterior neck is fully discussed in our preceding work [24]. The technical challenge of segmenting muscles is already solved [24] although as shown in supplementary material (SM) this paper demonstrates considerable improvement since our previously published work [24]. Here, we apply the deep learning methods developed in our lab and we test a series of hypothesis concerning the value of neck muscle boundaries for understanding dystonia (Fig. 2). By comparison with classification of sex and

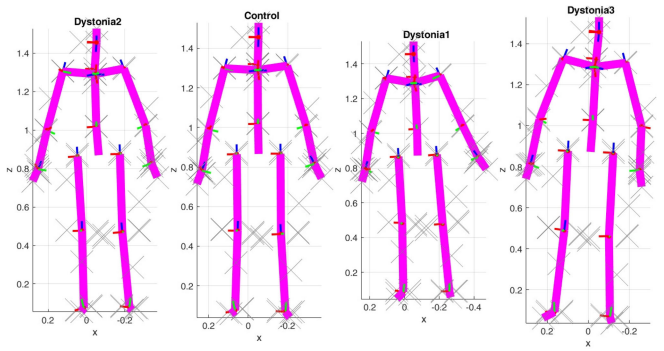

Fig. 10. **Median posture Dystonia sub-groups facing the reader**. Shows joint angles, median from all linked US-Vicon trials for each participant, averaged across all participants in group. This head referenced presentation shows the kinematic chain reconstructed from the head segment which is presented vertical and forward looking for each group. The red lines show the axis of extension and typically points to the right of the participant. The green lines show the axis of frontal rotation and typically points forwards. The blue lines show the axis of axial rotation and point along the segments long axis which is typically vertical. For the head, red, green blue axes indicate pitch, roll and yaw. For the trunk, red green and blue axes indicate forward lean, rightwards lean and right turn. Dystonia2 show neck deviated to their right relative to head, trunk deviated to their left relative to neck and back to the midline. Dystonia1 shows neck deviated to their left relative to head, trunk also deviated to their left. Dystonia 3 shows neck deviated to their right relative to head and trunk deviated further to their right.

age, our results demonstrate that US images of the neck muscles contain the information necessary to visualize, understand and potentially diagnose cervical dystonia (Table III, Figs. 7–11).

Having tested segmentation and classification of Dystonia using 60-fold cross validation ('Testing Mode') we sought to maximize the accuracy of analysis and visualization of dystonia. Our switch to use of 'Analysis Mode', for description and explanation of the predicted boundaries is justified by the purpose of the analysis. The purpose of cross-validation is to answer the question "how well do you expect your system to perform out in the real world on unseen data?" We have answered that with 60-Fold cross-validation. The purpose of our hypotheses is to see in optimal circumstances how dystonia manifests in ultrasound and how dystonia manifests in cross-sectional shapes of muscles. For that we used clinical labels to select (using SVM) the best training Mode. The clinical labels were not used in the process of annotation, training, validating and testing the neural networks. In practice, analysis and visualization of Testing Mode gave similar results to those presented, but the quality of the reconstruction of groups was lower ($\sim$72% rather than 85%, meaning the description was less accurate or complete.

C. Scientific and Clinical Value of Results

Cervical dystonia is a neurological disorder of sensorimotor integration characterized by abnormal postures of the head and neck. Abnormal involuntary dystonic activation of neck muscles is a primary symptom, but is a direct cause of pain, abnormal whole-body posture, and constraints on movement. Neck muscles traverse the primary link between the head (which is

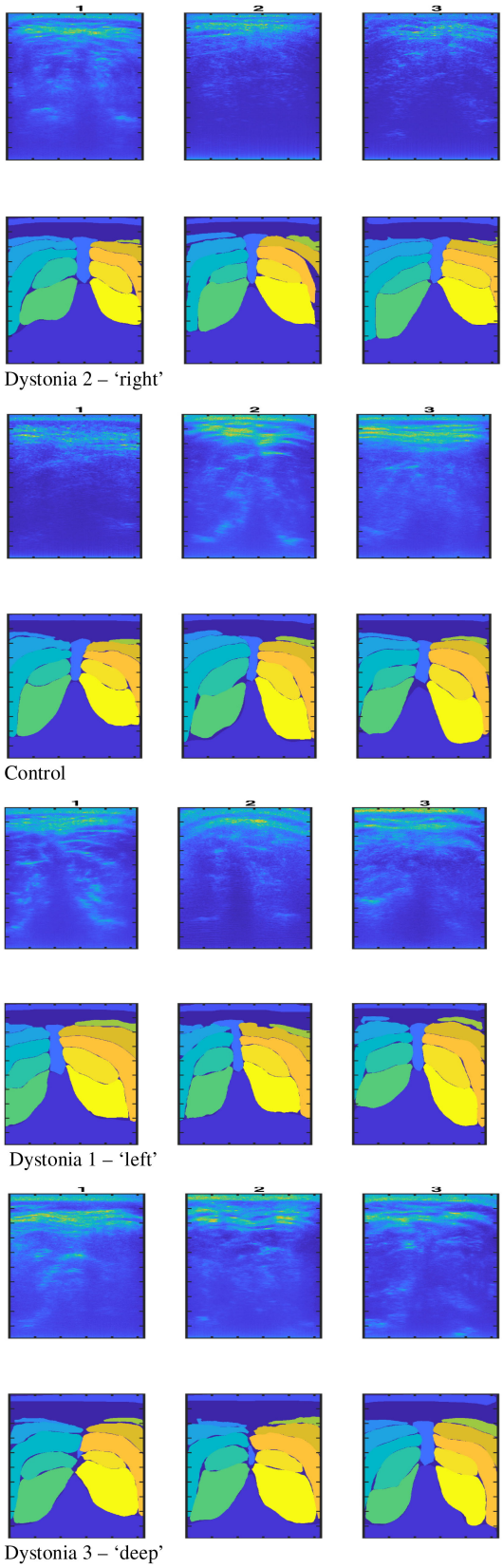

Fig. 11. **Examples of Dystonia sub-groups**. Each block contains the three example cases closest to the group center using discriminant function scores from Fig. 8. Each case shows the US image (above) and predicted labels (below).

the source of visual-vestibular head referenced sensory frames, location of sensory integration and motor planning) and the mass distribution of the body (trunk, upper and lower limbs). Abnormal action of neck muscles causes local changes in head and shoulder position and to maintain vertically aligned balance, these local changes require compensatory changes in whole body posture of the trunk and limbs. Neck muscles provide sophisticated proprioceptive sensation and have a primary role in integration of head-referenced with ground referenced coordinate frames which is also subject to interference by abnormal neck muscle activity. Altered body posture and sensory feedback is a consequence of abnormal neck muscle action. Thus, analysis of the neck muscles provides direct insight into cervical dystonia.

US imaging analysis can quantify dystonic muscle attributes (Figs. 8, 9, 10). US does not require participants with movement disorders to remain still and avoids limitations of MRI. US is relatively low cost and available in clinics.

The confidence measure provided by this neural network analysis (Fig. 4) gives inexperienced operators feedback to improve the quality of their US probe location and image. This analysis within clinic could facilitate communication between patient and clinician and would inform patients about their neck muscles and their specific dystonia. The objective recording of images and analysis provides a potential tool for guiding and recording the location of injections, for monitoring change and improvement with treatment, and thus is expected to improve the patient experience. In addition, our findings reinforce the potential critical role in CD of deep neck muscles, which have previously not been amenable to assessment or treatment.

D. Relationship to Previous Work

The application of machine learning and specifically deep learning to analysis of ultrasound images of muscle is rare [20], [22], [26]. While under-developed, the domain of muscle diagnosis is valuable since unlike visual observation, manual palpation or surface electromyography, ultrasound can see muscles deep within the body.

Segmentation is the foundation of muscle-specific analysis and recent methods providing segmentation of the neck muscles include computer vision [23] and deep learning approaches [24]. As shown in Supplementary Material, this study applies the most recent deep learning methods for this application. Following [24], this study uses direct manual annotation of US images to provide training labels. This approach allows us to develop training datasets for participants with a movement disorder who cannot remain still in an MRI machine [24]. Using metrics of JI and HD (Table I), the accuracy of segmentation achieved is consistent with existing benchmarks [23], [24]. The metric MHD (Table I) shows boundaries are typically accurate 1.9 ± 1.8 mm, and for deep muscles this accuracy allows for injection at average precision more than 90% for target sizes of several millimeter (SM Fig. A) and a margin from the muscle boundary of 3.7 ± 1 mm (Table II). These findings have important clinical implications, as freehand injections of botulinum toxin have been shown to have potentially suboptimal accuracy [27].

Prior to this study, it was an open question whether information contained within images of the neck muscles was of any value for diagnosis and understanding of cervical dystonia. This study affirms the US information content with respect to objective clinical labels (control, dystonia) and with respect to motion analysis.

Manual annotations provide only an approximation to the true muscle boundaries. With training, neural networks should learn image features that correlate consistently with the labels. In principle, machine learning should discard random error in human labels, and converge to the on-average correct answer within and between labelers.

The exciting result reported here (Table III, Fig. 6) is that these neural networks out performed their supervisors: we urge the reader to study SM Fig. 6 which shows very nice examples. These results confirm that neural networks encode information in the data consistent with supervisory labels, but which reduces the random error and non-generalizable component of the labels. These results (Table III) validate altogether, the method of labelling, the method of segmentation and the information content of ultrasound muscle images regarding dystonia.

This study demonstrates proof of concept of the feasibility of US imaging analysis of the neck muscles for understanding and diagnosing cervical dystonia. Figure 11 shows examples of dystonia sub-group categorization on the basis of eigen function scores and illustrate the immediate diagnosis that could be possible in the clinic. This proof of concept motivates further development of US technology. If deployed widely in clinics, there is potential to collect large quantities of data from the estimated 18,000 adults with this condition in the UK [1]. Combined with further exploration of neural network methods, there is potential for this tool to become very robust, and for a new domain (automated ultrasound muscle analysis) to be established. Evaluation of the effect of therapeutic interventions e.g., BoNT on the patterns of change on US would also be critical to determine the utility of this tool to monitor changes in dystonia severity, and to evaluate its utility as a potential biomarker.

E. Limitations

The current work contains several limitations. First, we have a relatively small number of cases, which may not encompass the full and expanding spectrum of neck movements seen in cervical dystonia [10] However we contribute our data to address the shortage of publically available examples (DOI: 10.23634/MMUDR.00624643). Further validation in a larger and independent clinical cohort would be desirable. Using more data, a clinical classifier would most logically be embedded within the neural network architecture. Second, this analysis is limited to an axial image at level C4. A larger number of probe locations/orientations and muscle images would be desirable. Third, this work predicts and interprets muscle shape, excluding prediction of texture and muscle activity. Further work will exploit the ultrasound information content revealing muscle function [20], [21] as well as geometry for a larger range of probe locations and orientations. We would expect segmentation

accuracy, boundary analysis and classification of clinical condition, all to depend to some extent on the quality of the image. We recommend to those readers replicating and extending our work, to choose an ultrasound machine, probe and machine settings providing the best possible image quality at full depth down to the vertebra.

V. Conclusion

This study provides the first application of deep learning to US imaging of the neck muscles in cervical dystonia and provides an automated objective visualization (c.f. Video) and subsequent pattern analysis of neck muscle boundaries. These results demonstrate that muscle boundaries extracted from a single axial image of the neck muscles have the information content to discriminate cervical dystonia from healthy controls and to visualize and understand the dystonic pattern of neck muscles. This proof of concept demonstrates potential for a clinical tool to provide objective online diagnosis of cervical dystonia, guidance and objective logging of injection sites, and objective monitoring of the effect of treatment with minimal requirement for operator expertise and minimal burden on clinical time. This work supports a case for further evaluation of an automated US-based tool in a larger longitudinal dataset.

Acknowledgment

With appreciation we acknowledge the contribution of all our participants who willingly gave their time, the support of Dr. Gregory May and Brian Bate to assist data collection and the encouragement and support of Joanne Day to develop this work to assist those with cervical dystonia.

References

[1] G. Defazio, J. Jankovic, J. L. Giel, and S. Papapetropoulos, "Descriptive epidemiology of cervical dystonia," *Tremor Other Hyperkinet. Mov. (N. Y).*, 2013, doi: 10.7916/D80C4TGJ.

[2] E. Tiderington *et al.*, "How long does it take to diagnose cervical dystonia?," *J. Neurol. Sci.*, vol. 335, no. 1–2, pp. 72–74, 2013.

[3] A. Albanese *et al.*, "Practical guidance for CD management involving treatment of botulinum toxin: A consensus statement," *J. Neurol.*, vol. 262, no. 10, pp. 2201–2213, 2015.

[4] G. Reichel, "Cervical dystonia: A new phenomenological classification for botulinum toxin therapy," *Basal Ganglia*, vol. 1, no. 1, pp. 5–12, 2011.

[5] H. A. Jinnah, E. Goodmann, A. R. Rosen, M. Evatt, A. Freeman, and S. Factor, "Botulinum toxin treatment failures in cervical dystonia: Causes, management, and outcomes," *J. Neurol.*, vol. 263, no. 6, pp. 1188–1194, Jun. 2016.

[6] W. H. Jost, H. Hefter, A. Stenner, and G. Reichel, "Rating scales for cervical dystonia: A critical evaluation of tools for outcome assessment of botulinum toxin therapy," *J. Neural Transm.*, vol. 120, no. 3, pp. 487–496, 2013.

[7] M. F. Contarino, M. Smit, J. van den Dool, J. Volkmann, and M. A. J. Tijssen, "Unmet needs in the management of cervical dystonia," *Front. Neurol.*, vol. 7, p. 165, 2016, doi: 10.3389/fneur.2016.00165.

[8] J. A. Van Gerpen, J. Y. Matsumoto, J. E. Ahlskog, D. M. Maraganore, and P. G. McManis, "Utility of an EMG mapping study in treating cervical dystonia," *Muscle and Nerve*, vol. 23, no. 11, pp. 1752–1756, 2000.

[9] A. Schramm, D. Huber, C. Mobius, A. Munchau, Z. Kohl, and T. Baumer, "Involvement of obliquus capitis inferior muscle in dystonic head tremor," *Parkinsonism Relat. Disord.*, vol. 44, pp. 119–123, Nov. 2017.

[10] L. Tatu and W. H. Jost, "Anatomy and cervical dystonia : 'Dysfunction follows form'," *J. Neural Transmiss.*, vol. 124, no. 2, pp. 237–243, Feb. 2017.

[11] C. Lungu, A. W. Tarulli, D. Tarsy, P. Mongiovi, V. G. Vanderhorst, and S. B. Rutkove, "Quantifying muscle asymmetries in cervical dystonia with electrical impedance: A preliminary assessment," *Clin. Neurophysiol.*, vol. 122, no. 5, pp. 1027–1031, May 2011.

[12] A. Schramm, T. Bäumer, U. Fietzek, S. Heitmann, U. Walter, and W. H. Jost, "Relevance of sonography for botulinum toxin treatment of cervical dystonia: An expert statement," *J. Neural Transmiss.*, vol. 122, no. 10, pp. 1457–1463, 2015.

[13] A. I. Grigoriu, M. Dinomais, O. Rémy-Néris, and S. Brochard, "Impact of injection-guiding techniques on the effectiveness of botulinum toxin for the treatment of focal spasticity and dystonia: A systematic review," *Arch. Physical Med. Rehabil.*, vol. 96, no. 11, pp. 2067–2078.e1, 2015.

[14] I. D. Loram, C. N. Maganaris, and M. Lakie, "Use of ultrasound to make noninvasive in vivo measurement of continuous changes in human muscle contractile length," *J. Appl. Physiol.*, vol. 100, no. 4, pp. 1311–1323, 2005.

[15] I. D. Loram, B. Bate, P. Harding, R. Cunningham, and A. Loram, "Proactive selective inhibition targeted at the neck muscles: This proximal constraint facilitates learning and regulates global control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 4, pp. 357–369, Apr. 2017.

[16] K. Bibbings, P. J. Harding, I. D. Loram, N. Combes, and E. F. Hodson-Tole, "Foreground detection analysis of ultrasound image sequences identifies markers of motor neurone disease across diagnostically relevant skeletal muscles," *Ultrasound Med. Biol.*, vol. 45, no. 5, pp. 1164–1175, Mar. 2019.

[17] I. D. Loram, C. N. Maganaris, and M. Lakie, "The passive, human calf muscles in relation to standing: The non-linear decrease from short range to long range stiffness," *J. Physiol.*, vol. 584, no. 2, pp. 661–675, 2007.

[18] P. J. Harding, I. D. Loram, N. Combes, and E. F. Hodson-Tole, "Ultrasound based detection of fasciculations in healthy and diseased muscles," *IEEE Trans. Biomed. Eng.*,vol. 63, no. 3, pp. 512–518, Mar. 2016.

[19] I. D. Loram, C. N. Maganaris, and M. Lakie, "Active, non-spring-like muscle movements in human postural sway: How might paradoxical changes in muscle length be produced?," *J. Physiol.*, vol. 564, no. 1, pp. 281–293, 2005.

[20] R. Cunningham, M. Sánchez, G. May, and I. Loram, "Estimating full regional skeletal muscle fibre orientation from B-Mode ultrasound images using convolutional, residual, and deconvolutional neural networks," *J. Imag.*, vol. 4, no. 2, p. 29, Jan. 2018.

[21] R. J. Cunningham, P. J. Harding, and I. D. Loram, "The application of deep convolutional neural networks to ultrasound for modelling of dynamic states within human skeletal muscle," 2017, *arXiv preprint arXiv:1706.09450*.

[22] R. J. Cunningham, P. Harding, and I. D. Loram, "Deep residual networks for quantification of muscle fiber orientation and curvature from ultrasound images," in *Communications in Computer and Information Science*, vol. 723, Cham: Springer, 2017, pp. 63–73.

[23] R. J. Cunningham, P. J. Harding, and I. D. Loram, "Real-time ultrasound segmentation, analysis and visualisation of deep cervical muscle structure," *IEEE Trans. Med. Imag.*, vol. 36, no. 2, pp. 653–665, Feb. 2017.

[24] R. J. Cunningham, M. B. Sánchez, and I. D. Loram, "Ultrasound segmentation of cervical muscle during head motion: A dataset and a benchmark using deconvolutional neural networks," engrXiv, Feb. 26, 2019, doi: 10.31224/osf.io/fsa3c.

[25] M. P. Dubuisson and A. K. Jain, "A modified Hausdorff distance for object matching," in *Proc. 12th Int. Conf. Pattern Recognit.*, 1994, vol. 1, pp. 566–568.

[26] P. Burlina, S. Billings, N. Joshi, and J. Albayda, "Automated diagnosis of myositis from muscle ultrasound: Exploring the use of machine learning and deep learning methods," *PLoS One*, vol. 12, no. 8, 2017, Art. no. e0184059.

[27] A. Schnitzler *et al.*, "Manual needle placement: Accuracy of botulinum toxin a injections," *Muscle Nerve*, vol. 46, no. 4, pp. 531–534, Oct. 2012.